Reduce RAM usage for neural network verification tool

Keywords: PyRAT, Neural Network, Graphical User Interface, Abstract Interpretation

Institution

The French Alternative Energies and Atomic Energy Commission (CEA) is a key player in research, development, and innovation. Drawing on the widely acknowledged expertise gained by its 16,000 staff spanned over 9 research centers with a budget of 4.1 billion Euros, CEA actively participates in more than 400 European collaborative projects with a large number of academic (notably as a member of Paris-Saclay University) and industrial partners. Within the CEA Technological Research Division, the CEA List institute addresses the challenges coming from smart digital systems.

Among other activities, CEA List's Software Safety and Security Laboratory (LSL) research teams design and implement automated analysis in order to make software systems more trustworthy, to exhaustively detect their vulnerabilities, to guarantee conformity to their specifications, and to accelerate their certification. Recently, the lab extended its activities on the topic of AI trustworthiness and gave birth to a new research group on the topic: AISER (Artificial Intelligence Safety, Explainability and Robustness). Developping, tools such as PyRAT, AIMOS or CAISAR to improve the safety of AI systems as a whole.

Objectives

Through the recent developments of AI, their safety has became a crucial concern in order to be able to use such AIs in critical systems. As such the AISER team has developed a neural network verification tool, PyRAT, which leverages the principles of abstract interpretation to verify neural networks. In comparison to classical software verification, PyRAT works directly on the weights, biases, and parameters of a neural network model thus making PyRAT lighter and faster to use for neural network analyses. PyRAT is developed in Python as it is a widely used language for neural network frameworks such as Keras, Pytorch or Tensorflow. As of now, the primary use of PyRAT is to assess robustness w.r.t. some perturbation around inputs on small to medium neural networks. However, a recent issue that arose when analysing large networks on image classification problems is that PyRAT's analysis is quite expensive in computation and RAM which could for such cases limit the usage of PyRAT to huge infrastructures.

The aim of this internship is thus to reduce the RAM footprint of a PyRAT analysis while keeping the same precision for the analysis. At the moment, it is possible to vastly reduce the amount of RAM used by simply removing precision, however, this would impact our capacity to prove anything on larger networks. The perspectives here are on two aspects: the sparsity of the matrices representing the noise of the analysis and the linearity of the operation. The first aspect means that we can try to use different data structures that a simple matrix to represent our introduced noise and we could thus gain in computation time and RAM usage. Indeed, as these matrix are very sparse before encountering a Fully Connected layer, we could envision using structures such as scipy sparse matrices to represent them. The second aspect that should be explored in this internship is that our matrices for the noise can also be separated in smaller matrices and calculated independently before re-joining them when needed. Here, this means we could do some part of the analysis then store it on the hard drive before repeating that for all other subpart before joining everything. This would allow to only use at one time a subpart worth of RAM at the cost of multiple disk writing and reading.

Qualifications

- **Minimal**
  - Master student or 2nd or 3rd year of engineering school
  - knowledge of Python
  - notions of AI and neural networks, and of AI frameworks (TF, Keras, Pytorch, ...)
  - ability to work in a team
- **Preferred**
  - some knowledge of abstract interpretation or formal method

Characteristics

- **Duration:** 5 to 6 months from september 2023
- **Location:** CEA Nano-INNOV, Paris-Saclay Campus, France

- **Compensation:**
  - €700 to €1300 monthly stipend (determined by CEA compensation grids)
  - maximum €229 housing and travel expense monthly allowance (in case a relocation is needed)
  - CEA buses in Paris region and 75% refund of transit pass
  - subsidized lunches

Application

If you are interested in this internship, please send to the contact persons an application containing:

- your resume;
- a cover letter indicating how your curriculum and experience match the qualifications expected and how you would plan to contribute to the project;
- your bachelor and master 1 transcripts;
- the contact details of two persons (at least one academic) who can be contacted to provide references.

Applications are welcomed until the position is filled. Please note that the administrative processing may take up to 3 months.

Contact persons

For further information or details about the internship before applying, please contact:

- Augustin Lemesle (augustin.lemesle@cea.fr)
- Zakaria Chihani (zakaria.chihani@cea.fr)